Appendix

Appendix A1.1 Study characteristics: Kirby, 2006, October (randomized controlled trial)

Characteristic	Description
Study citation	Kirby, P. C. (2006, October). I CAN Learn® in Orleans Parish Public Schools effects on LEAP 8th grade math achievement, 2003–2004. (Available from ed-cet, Inc. 2301 Killdeer Street, New Orleans, LA 70122.) *Additional source:* Kirby, P. C. (2004b, November). I CAN Learn® in Orleans Parish Public Schools effects on LEAP 8th grade math achievement, 2003–2004. (Available from ed-cet, Inc., 2301 Killdeer Street, New Orleans, LA 70122.)
Participants	Participants were 2,400 eighth-grade students (1,082 *I CAN Learn®* students and 1,318 traditional mathematics students) from 57 math classes in 13 Orleans Parish Public Schools.[1] Students were randomly assigned to intervention and comparison classrooms using the SASI Basic Scheduling software. Teachers were not randomized to conditions. Only students with no special education classification were included in the analyses. About 96% of the students in the sample were African-American, less than 1% were Hispanic, and less than 0.5% were Caucasian.[2]
Setting	The study took place in the Orleans Parish public school system, which includes the city of New Orleans. The participating schools were selected for this study based on two criteria: the schools included both traditional and *I CAN Learn®* eighth-grade classes and those classes included at least 20 students each.
Intervention	Students were taught using the *I CAN Learn®* mathematics curriculum. The amount of time students used the *I CAN Learn®* system varied by school, with some students using only the test prep module and others completing up to 95 lessons. The average number of lessons completed was 12.1.
Comparison	Comparison students were taught in traditional classes with the teacher serving as the primary deliverer of instruction. The author does not provide further information on the curriculum.
Primary outcomes and measurement	The primary outcome measure was the Louisiana Educational Assessment Program (LEAP) Grade 8 Mathematics Exam. (See Appendix A2 for more detailed descriptions of outcome measures.)
Teacher training	Information on teacher training was not provided.

1. The original sample included 14 schools. However, random assignment occurred in only 13 schools. For the purposes of this review, the study design and findings are reported for the subsample of 13 schools. Further, the study author conducted analyses of subsamples of students in addition to the total sample of 2,400 students in 13 schools. Those analyses focused on students of teachers who taught both conditions and students with matched pretest and posttest scores. For rating purposes, only the analysis of the total sample in the 13 schools was reviewed because, under the current WWC guidance, there is no attrition or confounding problem that might warrant a review of a subsample. In addition, the two additional analyses did not meet WWC evidence screens because of insufficient information on sample size and pretest equivalence.
2. The study notes that these data were collected in the Orleans Parish Public Schools before Hurricane Katrina. The demographics in the school district have changed since the study was conducted.

Appendix A1.2 Study characteristics: Kerstyn, 2001 (quasi-experimental design)

Characteristic	Description
Study citation	Kerstyn, C. (2001). Evaluation of the I CAN Learn® mathematics classroom: First year of implementation (2000–2001 school year). (Available from the Division of Instruction, Hillsborough County Public Schools, 901 East Kennedy Blvd., Tampa, FL 33602.)
Participants	Participants were 2,539 eighth-grade students in 116 classrooms (58 *I CAN Learn*® classrooms with 1,225 students and 58 traditional classrooms with 1,314 students) in Title I middle schools. The study was limited to regular education students. The students were racially diverse and many were eligible for free or reduced-price lunches. The 58 *I CAN Learn*® classrooms and 58 traditional classrooms were used in the analysis.
Setting	The participating students were from middle schools in the Hillsborough County Public School system in Florida. This county includes the Tampa metro area.
Intervention	Students were taught using the *I CAN Learn*® mathematics curriculum. The author does not indicate how many of the lessons are required to be completed for the curriculum to be implemented as intended. The author indicates that the *I CAN Learn*® system was implemented in 45-, 50-, 80-, and 90-minute class periods. When surveyed, the teachers reported that 45 minutes was not long enough to make it through the curriculum. The intervention and evaluation occurred during the 2000–01 academic year.
Comparison	Students in classrooms that were selected as a match with intervention classrooms were taught in traditional classes using a traditional math curriculum. The author does not provide further information on the curriculum.
Primary outcomes and measurement	The primary outcome measure was the Florida Comprehensive Assessment Test (FCAT) Grade 8 Math Test.[1] (See Appendix A2 for more detailed descriptions of outcome measures.)
Teacher training	Teachers in this study participated in training sessions on the use of the software and hardware, but not on use of the software in instruction.

1. A Semester 1 exam was also administered to study participants at the end of the first semester. The results are not reported in Appendix A3 with the results of the FCAT, because the psychometric properties of the Semester 1 exam were not reported and it measured the effect of the intervention at an earlier point than did the FCAT. Because the FCAT measured math achievement at the end of the school year, the WWC reasoned that the FCAT was a more appropriate and valid measure of the intervention's end-of-year effects.

Appendix A1.3 Study characteristics: Kerstyn, 2002, October (quasi-experimental design)

Characteristic	Description
Study citation	Kerstyn, C. (2002, October). Evaluation of the I CAN Learn® mathematics classroom: Second year of implementation (2001–2002 school year). (Available from the Division of Instruction, Hillsborough County Public Schools, 901 East Kennedy Blvd., Tampa, FL 33602.)
Participants	Participants were 11,125 eighth-grade students in 597 classes (129 *I CAN Learn*® classes with 1,871 students and 468 traditional mathematics classes with 9,254 students) enrolled in Algebra I, Algebra I Honors, MJ-3 pre-algebra, or MJ-3 Advanced math classes in the 41 middle schools in Hillsborough County Schools, Florida. The analyses incorporate all students. Subgroup analyses present separate results for standard curriculum students and FCAT-exempt students. Compared with traditional classrooms, a higher proportion of *I CAN Learn*® students were on free and reduced-price lunch and were from minority backgrounds (African-American and Hispanic).
Setting	Participating students were from middle schools in the Hillsborough County Public School system in Florida. This county includes the Tampa metro area.
Intervention	Students were taught using the *I CAN Learn*® *Algebra* course, which consists of 109 complete algebra lessons. The *I CAN Learn*® system is intended to be the primary source of instruction. The intervention and evaluation occurred during the 2001–02 academic year.
Comparison	Comparison students were taught using a traditional math curriculum. The author does not provide further information on the curriculum.
Primary outcomes and measurement	The primary outcome measure was the FCAT Grade 8 Math Test. (See Appendix A2 for more detailed descriptions of outcome measures.)
Teacher training	Information on teacher training was not provided.

Appendix A1.4 Study characteristics: Kirby, 2004, September (randomized controlled trial with teacher-intervention confound problem)

Characteristic	Description
Study citation	Kirby, P. C. (2004, September). Comparison of I CAN Learn® and traditionally-taught 8th grade general math student performance on the California Standards Test, Spring 2004. (Available from ed-cet, Inc., 2301 Killdeer Street, New Orleans, LA 70122.)
Participants	Participants were 204 eighth-grade students (91 *I CAN Learn®* students and 113 traditional mathematics students) in Bret Harte Middle School. The *I CAN Learn®* classrooms contained a higher proportion of African-American students and a lower proportion of Hispanic and non-native English speaking students than the comparison classrooms.
Setting	The participating students were from Bret Harte Middle School, which is one of five middle schools in Hayward Unified School District in Alameda County, California. Hayward, southeast of San Francisco and south of Oakland, had a population of 144,633 and a mean household income of $51,177, according to the 2000 Census. The city is ethnically diverse, with Hispanics (34%), whites (29%), Asians (21%), and African-Americans (11%).
Intervention	Students were taught eighth-grade mathematics by one teacher using the *I CAN Learn®* mathematics curriculum. The *I CAN Learn®* system consists of 303 lessons from basic mathematics to advanced algebra concepts. Teachers choose the lessons that align to local curricular goals. The *I CAN Learn®* system is intended to be the primary source of instruction. The intervention and evaluation occurred during the 2003–04 academic year.
Comparison	Comparison students were taught in traditional classes, with the teacher as the primary deliverer of instruction, using a curriculum based on the state-adopted Glencoe pre-algebra textbook.
Primary outcomes and measurement	The primary outcome measure was the General Mathematics exam from the California Standards Test (CST). (See Appendix A2 for more detailed descriptions of outcome measures.)
Teacher training	Information on teacher training was not provided.

Appendix A1.5 Study characteristics: Kirby, 2004a, November (randomized controlled trial with teacher-intervention confound problem)

Characteristic	Description
Study citation	Kirby, P. C. (2004a, November). Comparison of I CAN Learn® and traditionally-taught 8th grade student performance on the Georgia Criterion-Referenced Competency Test. Unpublished manuscript.
Participants	Participants were 254 eighth-grade students (93 in *I CAN Learn®* classes and 161 in traditional classes). All participants were regular education students.
Setting	One middle school in the Gilmer County School District in northwestern Georgia.
Intervention	Students were taught by one teacher using the *I CAN Learn®* system. The intervention was implemented during the 2003-04 school year.
Comparison	Comparison students were taught in traditional classes, with the teacher serving as the primary deliverer of instruction. The author did not specify what curriculum was used.
Primary outcomes and measurement	The primary outcome measure was the math exam from the Georgia Criterion-Referenced Competency Test (GCRCT). (See Appendix A2 for more detailed descriptions of outcome measures.)
Teacher training	Only one *I CAN Learn®* teacher in this study was trained in the year prior to implementation.

Appendix A1.6 Study characteristics: Kirby, 2005, January (randomized controlled trial with teacher-intervention confound problem)

Characteristic	Description
Study citation	Kirby, P. C. (2005, January). I CAN Learn® Algebra I in Catoosa County, Georgia. (Available from ed-cet, Inc., 2301 Killdeer Street, New Orleans, LA 70122.)
Participants	Participants were 137 ninth-grade students (84 *I CAN Learn*® students and 53 traditional mathematics students) enrolled in Algebra I during the 2004 winter semester.
Setting	The participating students were from suburban high schools (Lakeview-Fort Ogelthorpe High School in Catoosa County, Georgia, and in the metropolitan area of Chattanooga, Tennessee). The student population is 95% white; 27% of the students were eligible for free or reduced-price school lunches.
Intervention	Students were taught Algebra I by one teacher using the *I Can Learn*® *Algebra* curriculum. The *I CAN Learn*® *Algebra* course consists of 177 lessons and a 23,550-question question bank in 17 units. The *I CAN Learn*® system is intended to be the primary source of instruction.
Comparison	Comparison students were taught in traditional classes, with the teacher serving as the primary deliverer of instruction. Their curriculum was based on the Georgia Quality Core Curriculum standards, using the Larson, Boswell, Kanold, and Stiff (2001) *Algebra I* textbook published by McDougal Littell.
Primary outcomes and measurement	The primary outcome measure was the Georgia Algebra 1 End-of-Course Test (EOCT).
Teacher training	*I CAN Learn*® classes were taught by a teacher trained in a two-day summer workshop on how to use the software and how to manage the instructional environment.

Appendix A2 Outcome measures in the math achievement domain

Outcome measure	Description
Florida Comprehensive Assessment Test (FCAT) Grade 8 Math Test	The FCAT math test is a standardized measure that includes items related to all five content strands of Florida's *Sunshine State Standards* for mathematics: number sense, concepts, and operations; measurement; geometry and spatial sense; algebraic thinking; and data analysis and probability (as cited in Kerstyn, 2001; Kerstyn, 2002 October). Test content at grade 8 is evenly divided among these five content strands. Students are given 160 minutes to take the exam, which includes multiple-choice items, gridded-response items, and performance tasks. Test results are reported as scale scores, which range from 100 to 500.
General Mathematics exam from the California Standards Test (CST)	The General Mathematics CST for grade 8 is based on the California Mathematics Standards for grades 6 and 7 (as cited in Kirby, 2004 September). The CST is administered to students in grades 8 and 9 who have not yet completed or are not enrolled in discipline-specific standards-based math courses, or who are enrolled in the first year of a multi-year Algebra I course. The CST consists of 65 multiple-choice questions. Test results are reported as scale scores, which range from 150 to 600.
Louisiana Educational Assessment Program (LEAP) Grade 8 Mathematics Exam	The LEAP Mathematics test for grade 8 includes a multiple-choice section and an open-ended section for more complex tasks (as cited in Kirby, 2006 October). The test is aligned to the state's Comprehensive Curriculum and Grade Level Expectations. Six strands of the Louisiana Mathematics Framework are represented in LEAP: number and number relations; algebra; measurement; geometry; data analysis, probability, and discrete math; and patterns, relations, and functions. Test results are reported as scale scores, which range from 100 to 500.
Georgia Algebra 1 End-of-Course Test (EOCT)	The state-mandated Algebra 1 EOCT is aligned with the Georgia Quality Core Curriculum standards (as cited in Kirby, 2005 January). As of Fall 2004, the EOCT score is factored into students' Algebra I course grade. The Algebra 1 EOCT contains 90 questions related to five content domains: algebraic fundamentals, operations on real numbers and algebraic expressions, solving equations and inequalities, functions and their graphs, and connections and applications. Students pass the EOCT by scoring 600 or above, and the figures reported are based on the dichotomous pass/fail measure. The study presents the proportion of students passing the test.
Georgia Criterion-Referenced Competency Test (GCRCT) Math Test	The GCRCT is designed to measure how well students acquire the skills and knowledge described in the Georgia Performance Standards and the Quality Core Curriculum (as cited in Kirby, 2004a November). The GCRCT Math Test contains 60 items in six areas: number sense and numeration; geometry and measurement; patterns, relationships, and algebra; statistics and probability; computation and estimation; and problem solving. Test results are reported as scale scores, which range from 150 to 450.

Appendix A3 Summary of study findings included in the rating for the math achievement domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Author's findings from the study: Mean outcome (standard deviation[2]) — *I CAN Learn®* group	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[3] (*I CAN Learn®* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Kirby, 2006, October (randomized controlled trial)[7]								
LEAP Math scale scores	Grade 8	13/2,400	295.30 (37.60)	278.82 (43.50)	16.48	0.35	Statistically significant	+14
Average[8] for math achievement (Kirby, 2006 October)						0.35	Statistically significant	+14
Kerstyn, 2001, Algebra 1 (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	16/350	351.38 (30.80)	344.60 (28.36)	6.92	0.23	ns	+9
Average[8] for math achievement (Kerstyn, 2001 Algebra 1)						0.23	ns	+9
Kerstyn, 2001, Algebra 1 Honors (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	16/336	372.99 (34.47)	373.73 (35.80)	–0.74	–0.02	ns	–1
Average[8] for math achievement (Kerstyn, 2001 Algebra 1 Honors)						–0.02	ns	–1
Kerstyn, 2001, MJ-3 pre-algebra (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	64/1,420	296.77 (31.19)	293.89 (38.09)	2.88	0.08	ns	+3
Average[8] for math achievement (Kerstyn, 2001 MJ-3 pre-algebra)						0.08	ns	+3
Kerstyn, 2001, MJ-3 Advanced (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	20/430	332.51 (31.19)	327.40 (29.60)	5.11	0.17	ns	+7
Average[8] for math achievement (Kerstyn, 2001 MJ-3 Advanced)						0.17	ns	+7
Kerstyn, 2002, October, Algebra 1 (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	66/1,195	347.40 (20.90)	349.80 (21.00)	–2.36	–0.11	ns	–5
Average[8] for math achievement (Kerstyn, 2002 October, Algebra 1)						–0.11	ns	–5

(continued)

Appendix A3 Summary of study findings included in the rating for the math achievement domain *(continued)*

Outcome measure	Study sample	Sample size (schools/ students)	Author's findings from the study: Mean outcome (standard deviation[2]) — *I CAN Learn®* group	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[3] (*I CAN Learn®* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Kerstyn, 2002, October, Algebra 1 Honors (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	91/1,894	369.60 (27.80)	374.30 (27.60)	–4.73	–0.17	ns	–7
Average[8] for math achievement (Kerstyn, 2002 October, Algebra 1 Honors)						–0.17	ns	–7
Kerstyn, 2002, October, MJ-3 pre-algebra (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	328/5,957	293.80 (32.00)	289.90 (40.50)	3.91	0.10	Statistically significant	+4
Average[8] for math achievement (Kerstyn, 2002 October, MJ-3 pre-algebra)						0.10	Statistically significant	+4
Kerstyn, 2002, October, MJ-3 Advanced (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	112/2,079	329.90 (23.90)	331.00 (24.10)	–1.11	–0.05	ns	–2
Average[8] for math achievement (Kerstyn, 2002 October, MJ-3 Advanced)						–0.05	ns	–2
Kirby, 2004, September (randomized controlled trial with teacher-intervention confound problem)[7]								
General Mathematics CST	Grade 8	1/204	315.60 (45.80)	299.70 (49.70)	15.85	0.33	Statistically significant	+13
Average[8] for math achievement (Kirby, 2004 September)						0.33	Statistically significant	+13
Kirby, 2004a, November (randomized controlled trial with teacher-intervention confound problem)[7]								
GCRCT[9]	Grade 8	1/254	333.50 (35.70)	319.90 (31.70)	13.60	0.41	Statistically significant	+16
Average[8] for math achievement (Kirby, 2004a November)						0.41	Statistically significant	+16

(continued)

Appendix A3 Summary of study findings included in the rating for the math achievement domain *(continued)*

Outcome measure	Study sample	Sample size (schools/ students)	Author's findings from the study: Mean outcome (standard deviation[2]) — *I CAN Learn®* group	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[3] (*I CAN Learn®* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Kirby, 2005, January (randomized controlled trial with teacher-intervention confound problem)[7]								
Algebra 1 EOC test	Grade 9	1/137	0.86 (0.35)	0.72 (0.45)	0.14	0.52	Statistically significant	+20
Average[8] for math achievement (Kirby, 2005 January)						0.52	Statistically significant	+20
Domain average[8] for math achievement across all studies						0.15	na	+6

ns = not statistically significant

na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices. Subgroup findings from the same studies are not included in these ratings but are reported in Appendix A4.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The mean difference is adjusted for pretest differences in Kerstyn (2002, October). For each sample reported in Kerstyn (2002, October), the control group mean is the intercept from the HLM model and the intervention group mean is this intercept plus the type of classroom (intervention or comparison) HLM beta coefficient.
4. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations. The effect size for Kirby (2005, January) is estimated based on the Cox Transformation (logs odds ratio divided by 1.65).
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
7. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Kerstyn (2001) and Kirby (2006), corrections for clustering were needed. No other studies required corrections for clustering or multiple comparisons.
8. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect size.
9. The author reported results from the Georgia Criterion-Referenced Competency Test as scale scores and as criterion scores (that is, the percentage that passed the criterion score compared with the percentage that failed), but the WWC focused on the results from the scale scores because they contain more information than categorical scores.

Appendix A4 Summary of subgroup findings for the math achievement domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Author's findings from the study: Mean outcome (standard deviation[2]) — *I CAN Learn®* group	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[3] (*I CAN Learn®* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Kerstyn, 2002, October, MJ-3 pre-algebra standard curriculum (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	258/4,045	294.20 (33.90)	295.10 (33.50)	–0.90	–0.03	ns	–1
Kerstyn, 2002, October, MJ-3 pre-algebra FCAT-exempt (quasi-experimental design)[7]								
FCAT mathematics	Grade 8	249/888	299.50 (37.60)	284.40 (52.90)	15.15	0.29	Statistically significant	+11

ns = not statistically significant

1. This appendix presents additional subgroup findings for measures that fall in the math achievement domain. The findings used for rating purposes are presented in Appendix A3.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
7. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Kerstyn (2002, October), no correction for clustering was needed.

Appendix A5 *I Can Learn*® rating for the math achievement domain

The WWC rates an intervention's effects in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of math achievement, the WWC rated *I CAN Learn*® *Pre-Algebra* and *Algebra* as having positive effects. The remaining ratings (potentially positive effects, mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *I CAN Learn*® was assigned the highest applicable rating.

Rating received
Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.
• Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.
Met. Five studies of *I CAN Learn*® showed statistically significant positive effects. Of those, one study had a strong design according to WWC standards.
• Criterion 1: No studies showing statistically significant or substantively important *negative* effects.
Met. Five studies of *I CAN Learn*® showed statistically significant positive effects. The remaining seven studies showed indeterminate effects. No studies showed statistically significant or substantively important negative effects.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain level effect. The WWC also considers the size of the domain level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.